USDA Service Center Agencies
Geospatial Data Management Team
Data Management Plan For

Official Soil Survey Theme
(Includes Soil Survey Geographic Database or SSURGO)

January 2005
Ken Harward and Jim Fortner

I. Purpose and Scope (business case)

A. Purpose

The Natural Resources Conservation Service (NRCS) has federal responsibility and leadership for the National Cooperative Soil Survey (NCSS). The primary focus of the NCSS has been mapping and classifying the soils of the U.S. and providing this information in the form of static, printed soil survey reports. This focus has shifted to maintaining and providing a dynamic resource of soil information that can serve diverse, individualized needs with digital products tailored to meet these needs.

Soil Data Warehouse/ Soil Data Mart – The Soil Data Warehouse (SDW) is a facility to house current and archived, time-stamped versions of official soil survey data and information for each soil survey area in the United States. Data in the SDW consists of digital soil map data (where available), tabular data describing individual soil map units and map unit components, interpretations of those data, and Federal Geographic Data Committee (FGDC) compliant metadata describing procedures used in the development of the data and the database in which the data are stored.

The Soil Data Mart (SDM) is a web interface used by clients and customers to access the current version of official data housed in the SDW. The user can generate standard reports of soil data or download the data to his/her local computer.

The purpose of the warehouse and data mart is to provide a single point of delivery of our official soil survey information, whether it is to the Field Office Technical Guide (FOTG), Soil Survey Geographic Database (SSURGO), Customer Service Toolkit, Technical Service Providers, or the general public.

B. Scope

The scope of the soil database includes the continental US, Caribbean Area, Alaska, Hawaii, and the Pacific Basin. The map extent for a Soil Survey Geographic (SSURGO) data set is a soil survey area, which may consist of a county, multiple counties, or parts of one or more counties. A SSURGO data set consists of digital map data, attribute data, and metadata. SSURGO data are available for selected counties and areas throughout the United States and its territories. For some survey areas, only attribute data and associated metadata are available.

II.   Acquisition

A. Data Source

1. Producer Information

a. Name

The data originates from National Cooperative Soil Survey (NCSS) field data collection efforts. The NCSS is a partnership of federal, state and local entities including NRCS, USFS, BLM, BIA, NPS, state universities and agencies, county and local units of government, and private consultants.

b. Location of Headquarters

Natural Resources Conservation Service
Soil Survey Division
Room 4250 South Building
14th & Independence Ave, SW
Washington, DC 20250

National Soil Survey Center
USDA - Natural Resources Conservation Service
Federal Building, Room 152
100 Centennial Mall, North
Lincoln, NE 68508-3866
Telephone: (402) 437-4000

NRCS State Offices (State Soil Scientist)

c. Internet Addresses

Main soil survey web site: http://soils.usda.gov
National Soil Survey Center: http://soils.usda.gov/contact/nssc/
Soil Data Mart: http://soildatamart.nrcs.usda.gov/
Geospatial Data Gateway: http://datagateway.nrcs.usda.gov/
State Soil Scientist: http://soils.usda.gov/contact/state_offices/ or
http://xx.nrcs.usda.gov (where xx is replaced with the two letter state abbreviation for the state of interest)

2. Publisher Information

a. Name

The State Soil Scientist in the state where a specific soil survey is located is the primary publisher of the data.

Soil data is available through the Soil Data Mart at http://soildatamart.nrcs.usda.gov/

b. Location of Headquarters

Natural Resources Conservation Service
Soil Survey Division
Room 4250 South Building
14th & Independence Ave, SW
Washington, DC 20250

National Soil Survey Center

USDA - Natural Resources Conservation Service
Federal Building, Room 152
100 Centennial Mall, North
Lincoln, NE 68508-3866
Telephone: (402) 437-4000

NRCS State Offices (State Soil Scientist)

c. Internet Address

Main soil survey web site: http://soils.usda.gov
National Soil Survey Center: http://soils.usda.gov/contact/nssc/
Soil Data Mart: http://soildatamart.nrcs.usda.gov/
Geospatial Data Gateway: http://datagateway.nrcs.usda.gov/
State Soil Scientist: http://soils.usda.gov/contact/state_offices/ or
http://xx.nrcs.usda.gov (where xx is replaced with the two letter state abbreviation for the state of interest)

3. Acquisition Information
a. Delivery Media

Soil survey data is available via FTP download through the Soil Data Mart or the Geospatial Data Gateway. It is also available on CD or DVD through the Geospatial Data Gateway.

b. Download URL

Soil Data Mart: http://soildatamart.nrcs.usda.gov/
Geospatial Data Gateway: http://datagateway.nrcs.usda.gov/

c. Data Availability

Available soil survey datasets are listed on the following web sites:
Soil Data Mart: http://soildatamart.nrcs.usda.gov/
Geospatial Data Gateway: http://datagateway.nrcs.usda.gov/

B. *Standards Information*
1. Geospatial Data Standard
a. Standard Name and Steward Information

National Standard for Soil Geographic Data
Soil Data Subcommittee
Federal Geographic Data Committee

b. Standard Version

November 1999

c. Standard URL

http://www.ftw.nrcs.usda.gov/ssurgo.txt (useful README on SSURGO)

2. Metadata Standard
a. Standard Name and Steward Information

Metadata are compliant with:
Federal Geographic Data Committee (FGDC)
Content Standard for Digital Geographic Metadata FGDC
STD-001-1998 Version 2 revised June 1998

And:
United States Department of Agriculture (USDA) Service Center Agencies (SCA)
Standard for Geospatial Dataset Metadata
SCI Std 003-02 October 15, 2003

http://www.itc.nrcs.usda.gov/scdm/docs/SPG-GeospatialDatasetFileMetadata.pdf

b. Description of Metadata Captured
A metadata text file is distributed with each soil survey data set. The metadata text file provides information on the content, quality, condition, and other characteristics of the data. The sections of metadata include the following:

> Data Set Identification Information
> Data Quality Information
> Spatial Reference Information
> Status Information
> Lineage (processing steps)
> Entity and Attribute Information
> Distribution Information
> Data Use Information
> Metadata Reference Information

c. Metadata Accuracy and Completeness Assessment

The metadata is typically complete. Additional information and a metadata template are available in the geospatial data standard.

C. Acquired Data Structure
1. Geospatial Data Format
a. Format (raster, vector, etc.)

The map data is vector. A SSURGO data set consists of map data, attribute data, and metadata. SSURGO data is collected in 7.5-minute quadrangle units, and distributed as complete coverage for a soil survey area. Computer software is used to join soil boundaries ending at quad neatlines to adjoining maps to achieve an exact match. Map data consists of polygons, lines, and / or points. (For more details, see http://soildatamart.nrcs.usda.gov/documents/SSURGODataPackagingandUse.pdf)

b. Format Name

SSURGO map data are available in ESRI shape file, Arc INFO coverage, and Arc interchange file formats. Attribute data are distributed in ASCII format. Metadata are in ASCII format.

c. Data Extent

SSURGO data are available for selected counties and areas throughout the United States and its territories. The map extent for a Soil Survey Geographic (SSURGO) data set is a soil survey area, which may consist of a county, multiple counties, or parts of one or more counties.

d. Horizontal and Vertical Resolution

The accuracy of these digital data is based upon their compilation to base maps that meet National Map Accuracy Standards. The difference in positional accuracy between the soil boundaries and special soil features locations in the field and their digitized map locations are unknown. The location accuracy of soil delineation on the ground varies with the transition between map units.

For example, on long, gently sloping landscapes, transition occurs gradually over many feet. Where landscapes change abruptly from steep to level, the transition will be very narrow. Soil delineation boundaries and special soil features generally were digitized within 0.01 inch of their locations on the digitizing source. The digital map elements are edge matched between data sets. The data along each quadrangle edge are matched against the data for the adjacent quadrangle. Edge locations generally do not deviate from centerline to centerline by more than 0.01 inch.

e. Absolute Horizontal and Vertical Accuracy

Spatial data meet NRCS standards and specifications for digitizing outlined in Section 647.07 Digitizing specifications in part 647 Soil Map Development of the NRCS National Soil Survey Handbook.

f. Nominal Scale

Maps are made at scales ranging from 1:12,000 to 1:63,360. Typically scales are 1:12,000 or 1:24,000.

g. Horizontal and Vertical Datum

The datum is North American Datum 1983 for all appropriate areas (UTM zone 3 through 22) and World Geodetic System 1984 elsewhere. The vertical datum is mean sea level.

h. Projection

Geographic, Universal Transverse Mercator, and State Plane for ESRI shapefile, Arc coverage, or Arc interchange

i. Coordinate Units

Coordinate units for survey area data vary with the requested projection. Data delivered in Geographic projection (NAD83 or WGS84, as appropriate) are in decimal degrees, data projected to UTM zone or State Plane systems are in meters.

Coordinate units for DLG are meters.

j. Average Data Set Size

For one Soil Survey Area:

Depending on spatial data format chosen, and size of geographic area, files average about 28 megabytes (compressed). The metadata file averages about 30 kilobytes.

k. Symbology

None

2. Attribute Data Format

    a.   Format Name

        ASCII files. A Microsoft Access database template is available from the soil data mart to convert ASCII files to database format.

    b.   Database Size

        About two megabytes of attribute data (compressed) per ASCII data set.

3. Data Model

    a.   Geospatial Data Structure

        ESRI shapefile, Arc coverage or Arc interchange including:
1. Soil Survey Area Boundary Polygon(s) (Required)
2. Map Unit Boundary Polygons (Required)
3. Line Map Units (Optional)
4. Point Map Units (Optional)
5. Line Spot Features (Optional)
6. Point Spot Features (Optional)
7. Spot Feature Descriptions (Required if Line Spot Features or Point Spot Features are included)

    b.   Attribute Data Structure

        ASCII tables for all soil survey data

    c.   Database Table Definition

        Information from:
        http://soildatamart.nrcs.usda.gov/SSURGOMetadata.aspx

    d.   Data Relationship Definition

        http://soildatamart.nrcs.usda.gov/documents/SSURGODataModel.pdf

e.   Data Dictionary

The following link references several files that describe the tables, attributes, indexes, relationships, and domains for the soil database.
http://soildatamart.nrcs.usda.gov/SSURGOMetadata.aspx

D. Policies
1. Restrictions
a.   Use Constraints

The U.S. Department of Agriculture, Natural Resources Conservation Service, should be acknowledged as the data source in products derived from these data.

This data set is not designed for use as a primary regulatory tool in permitting or siting decisions, but may be used as a reference source. This is public information and may be interpreted by organizations, agencies, units of government, or others based on needs; however, they are responsible for the appropriate application. Federal, State, or local regulatory bodies are not to reassign to the Natural Resources Conservation Service any authority for the decisions that they make. The Natural Resources Conservation Service will not perform any evaluations of these maps for purposes related solely to State or local regulatory programs.

Photographic or digital enlargement of these maps to scales greater than at which they were originally mapped can cause misinterpretation of the data. If enlarged, maps do not show the small areas of contrasting soils that could have been shown at a larger scale. These data and their interpretations are intended for planning purposes only. The depicted soil boundaries, interpretations, and analysis derived from them do not eliminate the need for onsite sampling, testing, and detailed study of specific sites for intensive uses. Digital data files are periodically updated. Files are versioned and dated, and users are responsible for obtaining the latest version of the data.

b.   Access Constraints

None

c.   Certification Issues

The state soil scientist has certified the data on the Soil Data Mart. SSURGO data has been certified that it meets required specifications.

2. Maintenance
a.   Temporal Information

Digital data files are periodically updated. Soil survey datasets are dated and versioned, and users are responsible for obtaining the latest version of the data.

b.   Average Update Cycle

Periodic, as needed

E. Acquisition Cost
1. Cooperative Agreement
a.   Description of Agreement

None

b. Status of Agreement

None

2. Cost to Acquire Data

FTP is free. $US 50 for CD-ROM, or $US 100 for DVD of SSURGO map and attribute data.

III. Integration

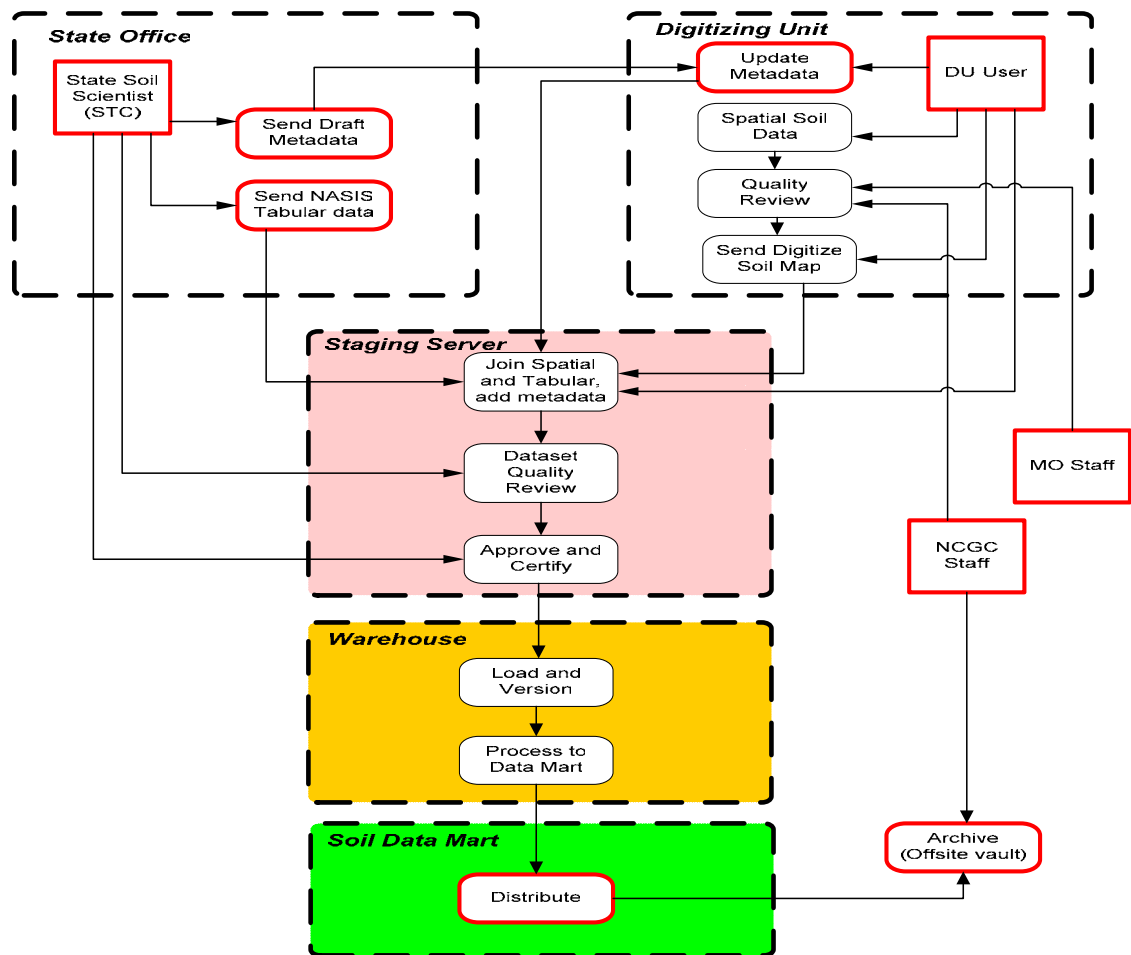## A. *Value Added Process*

1. Benefit to the Service Center

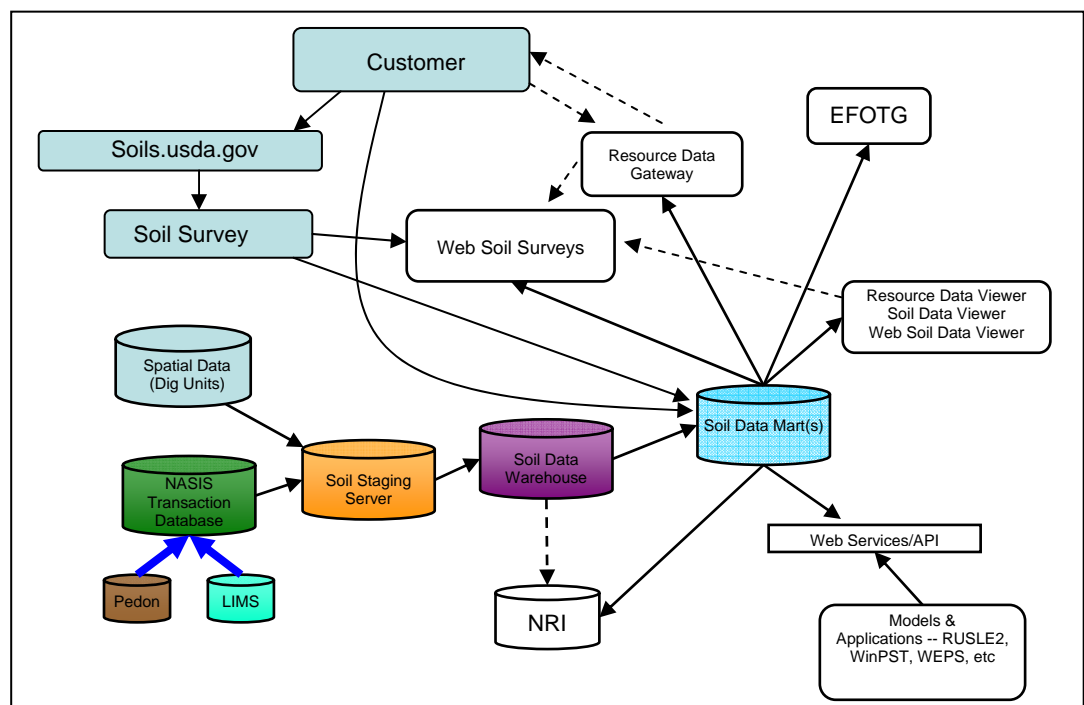Soil data is identified as a mission critical data layer for use within service center agencies.

2. Process Description

a) The following diagram shows the basic process flow from publisher to distribution point.

b) Process flow showing use by business applications, clients and customers.

3. Technical Issues

a. Tiling

A SSURGO data set is a soil survey area, which may consist of a county, multiple counties, or parts of multiple counties. This could be changed to a county tiling scheme. However, service center personnel are accustomed to one or more soil survey areas for the service center and perhaps prefer to work with multiple survey maps instead of one soil map for the service center. In addition, the system may be more responsive working with multiple soil survey area maps.

b. Compression

None

c. Scale

The publication scale for new and updated soil surveys is 1:12,000 or 1:24,000. (Alaska for 1:25,000 and the Caribbean for 1:20,000 are accepted.) Other scales require approval by the Director, Soil Survey Division, of the Natural Resources Conservation Service. (See http://soils.usda.gov/technical/handbook/contents/part647.html#01.) As a result, the data should not be used at scales larger than the publication scale.

d. Tonal Matching

This is not applicable to vector data, although soil maps are compiled on digital orthophotography that should be of high quality including tonal matching.

e. Edge-matching

A SSURGO data set is a soil survey area, which may consist of a county, multiple counties, or parts of multiple counties. Part of the compilation process requires edge matching between individual survey sheets and edge matching with all surrounding soil survey areas (http://soils.usda.gov/technical/handbook/contents/part647p3.html#ex3).

4. Quality Control

a. Procedures

Soil survey quality control is the process of providing direction, inspection, and coordination of soil survey activities to ensure that soil survey products meet the defined standards for content, accuracy, and precision. These are detailed in the National Soil Survey Handbook (NSSH) Part 609 at the following URL: http://soils.usda.gov/technical/handbook/contents/part609.html

b. Acceptance Criteria

The acceptance criteria are detailed in several parts of the NSSH including Parts 608, 609, 610, 644, 647, 648, and 651. See the following URL for more information: http://soils.usda.gov/technical/handbook/contents.html

5. Data Steward
    a. Name and Organization

        Currently, the data steward for the integrated data is:

        National Soil Survey Center
        USDA - Natural Resources Conservation Service
        100 Centennial Mall North
        Federal Building, Rm 153
        Lincoln, Nebraska 68508 USA
        Attn: Jim R. Fortner, Soil Scientist

    b. Responsibilities

        Store the SSURGO data and keep it accessible via ftp.

B. Integrated Data Structure
1. Geospatial Data Format
    a. Format (raster, vector, etc.)

        Vector

    b. Format Name

        ESRI Shape file, ArcInfo Coverage, ArcInfo Interchange

    c. Data Extent

        Soil survey area

    d. Horizontal and Vertical Resolution

        Same as source data

    e. Absolute Horizontal and Vertical Accuracy

        Same as source data

    f. Nominal Scale

        Same as source data

    g. Horizontal and Vertical Datum

        The horizontal datum is the North American Datum (NAD) 83. The vertical datum is mean sea level. The vertical datum is mean sea level.

    h. Projection

        Universal Transverse Mercator (UTM) in North American Datum (NAD) 83 or World Geodetic System 84.

    i. Coordinate Units

        Meters

j. Symbology
Symbols for point and line special features are selected from available ESRI palettes. There is a need for an NRCS point and line feature palette to be developed in agreement with the Conventional and Special Symbols Legend NRCS-SOILS-37A. See http://soils.usda.gov/technical/handbook/contents/part627p3.html#ex5 and http://www.itc.nrcs.usda.gov/scdm/docs/SPG-StandardforGeospatialSymbology.pdf

2. Attribute Data Format

a. Format Name

Dbase V, as part of an ESRI Shape file. Soil attribute data is available from the Soil Data Mart (http://soildatamart.nrcs.usda.gov) or through the Geospatial Data Gateway (http://datagateway.nrcs.usda.gov) as ASCII files. Scripts and templates are available to load the ASCII files into an MS Access database.

b. Database Size

The database size depends on the extent of the soil survey area and the variability of soils. It ranges from 10 MB to more than 100 MB.

3. Data Model

a. Geospatial Data Structure

Poly Files		Point Files		Line Files	
map shp	shp file	map shp	shp file	map shp	shp file
map dbf	dbf file	map dbf	dbf file	map dbf	dbf file
map shx	shx file	map shx	shx file	map shx	shx file
map sbn	sbn file	map sbn	sbn file	map sbn	sbn file
map sbx	sbx file	map sbx	sbx file	map sbx	sbx file

b. Attribute Data Structure

Dbase V, as part of an ESRI Shape file. Attribute data structure is located here: http://soildatamart.nrcs.usda.gov/SDMService.aspx

c. Database Table Definition

Standard .dbf file that goes with shape file. For polygon data, the .dbf file includes the additional field musym.

For point data, the .dbf file includes the additional fields featsym and symbol (in some soil survey areas).

For line special feature data, the .dbf file includes the additional field featsym

Attribute database table definitions are located here http://soildatamart.nrcs.usda.gov/SDMService.aspx.

d. Data Relationship Definition

Data relationship definitions are located here:
http://soildatamart.nrcs.usda.gov/SDMService.aspx.

e.   Data Dictionary

Attribute data dictionary is located here:
http://soildatamart.nrcs.usda.gov/SDMService.aspx.

C. Resource Requirements
1. Hardware and Software

To acquire and integrate one set of Soil Survey Area data, a Windows Pentium class machine with approximately 500 megabytes of free disk space is required. ESRI ArcView software is required for the spatial data. Microsoft Access is required for the tabular data.

2. Staffing

The access and integration of soil data is automated through web sites and downloadable scripts. Anyone with basic to intermediate to advanced computer skills can acquire soil survey data. Intermediate to advanced computer skills are required to integrate soil survey data with other datasets.

D. Integration Cost
1. Hardware and Software

To reformat, reproject, and subset the dataset a minimum the following is required:
ArcInfo on UNIX or Windows platform
ArcView on Windows platform
Five gigabyte disk

2. Staffing

Intermediate to advanced ESRI ArcInfo or ArcView skills and MS Access database skills are required for reformatting, reprojecting, and subsetting the soil database.

IV.   Delivery

A. Specifications
1. Directory Structure
a.   Folder Theme Data directory structure:
F:\Geodata
   \soil_<stssaid>
     \tabular
     \spatial

2. File Naming Convention

a. List of Theme Files and The File Naming Convention
http://www.itc.nrcs.usda.gov/scdm/docs/SPG-GeospatialDataSetFileNamingStandard.pdf

Subfolder Name	File Name Prefix	Description
\soil__<stssaid>\tabular	soil_d_<stssaid>	Access database of soil survey attribute data in the current SSURGO structure format. <stssaid> is the State Soil Survey Area ID (e.g., soil_d_ca048)
\soil__<stssaid>\spatial	soilsa_a_<stssaid>	Soil survey area boundary polygons
\soil__<stssaid>\spatial	soilmu_a_<stssaid>	Soil map unit boundary polygons
\soil__<stssaid>\spatial	soilmu_l_<stssaid>	Soil line map units
\soil__<stssaid>\spatial	soilmu_p_<stssaid>	Soil point map units
\soil__<stssaid>\spatial	soilsf_l_<stssaid>	Soil line spot features
\soil__<stssaid>\spatial	soilsf_p_<stssaid>	Soil point spot features
\soil__<stssaid>\spatial	soilsf_t_<stssaid>	Soil spot features description
\soil__<nnnnnn>\tabular	soilmosaic_d_<nnnnnn>	Merged attribute data for more than one soil survey area to support service center area of service. <nnnnnn > is the OIP office ID not OIP site ID
\soil__<nnnnnn>\spatial	soilmosaicsa_a_<nnnnnn>	Merged soil survey area boundary polygons for more than one soil survey area to support service center area of service.
\soil__<nnnnnn>\spatial	soilmosaicmu_a_<nnnnnn>	Merged soil map unit boundary polygons for more than one soil survey area to support service center area of service.
\soil__<nnnnnn>\spatial	soilmosaicmu_l_<nnnnnn>	Merged line map units for more than one soil survey area to support service center area of service.
\soil__<nnnnnn>\spatial	soilmosaicmu_p_<nnnnnn>	Merged point map units for more than one soil survey area to support service center area of service.
\soil__<nnnnnn>\spatial	soilmosaicsf_l_<nnnnnn>	Merged line spot features for more than one soil survey area to support service center area of service.
\soil__<nnnnnn>\spatial	soilmosaicsf_p_<nnnnnn>	Merged point spot features for more than one soil survey area to support service center area of service.
\soil__<nnnnnn>\spatial	soilmosaicsf_t_<nnnnnn>	Merged spot feature descriptions for more than one soil survey area to support service center area of service.

B. User Information

1. Accuracy Assessment

a. Alignment with Other Theme Geospatial Data

The data are captured at scales varying from 1:12,000 to 1:63,360, with the vast majority at 1:12,000 or 1:24,000. There should be general alignment with the ortho-photo layer. If other data layers are overlaid, alignment will likely not be perfect due to the fact that the data layers may have been captured at different scales.

b.   Content

The dataset contains soil maps, soil property attribute data, and soil interpretations. The soil property attribute data are based on measured values of representative soils which is then extrapolated over the geographic extent of a given map unit. This is the best known soils database available of its kind.

2. Appropriate Uses of the Geospatial Data
a.   Display Scale

The original data source scale or smaller, usually 1:12,000 or 1:24,000.

b.   Plot Scale

The original data source scale or smaller, usually 1:12,000 or 1:24,000.

c.   Area Calculations

Area Calculations are as accurate as the source data, capture scale and algorithm used by ArcInfo/ArcView.

d.   Decision Making

These data and their interpretations are intended for planning purposes only. The depicted soil boundaries, interpretations, and analysis derived from them do not eliminate the need for onsite sampling, testing, and detailed study of specific sites for intensive uses.

C. Maintenance and Updating
1. Recommendations and Guidelines
a.   Original data location and structure

The official dataset of soil survey data for each soil survey area is stored in the Soil Data Warehouse, and is available for distribution from the Soil Data Mart and the Geospatial Data Warehouse.   At SCA field offices a working copy will be placed in **F:\geodata\ soil_<stssaid>** folder using the standard naming convention identified above. The spatial data should be in ArcView shapefile format until other formats are needed. Tabular data are stored in MS Access database format.

b.   Update Cycle

There is no specified update cycle for these data. As the data are part of the Field Office Technical Guide, review is required at least every five years as per NRCS General Manual policy. Each state is responsible for maintenance of their respective datasets. A variety of factors influence their decision to update a particular dataset, including new agency or department program data needs and updated, more complete information. Tabular data can be updated independent of the associated spatial data. As an updated dataset is published to the Soil Data Warehouse, it is made available for distribution via the Soil Data Mart and Geospatial Data Warehouse. SCA field offices are responsible for ensuring that they have the most up-to-date version of the data available.

c.   Availability

Each state is responsible for publishing each of their soil survey area datasets to the Soil Data Warehouse and Soil Data Mart. Datasets are to be available for all completed soil survey areas. In addition to this, data for partially completed soil surveys may also be available. Tabular data will be available for each survey area. Digital spatial data will be available where digitized. These data are available for distribution to any user, including the general public.

d. Change control

Change control is related to both updated dataset content and changes to data structure of those datasets. Updates to data content may occur at various times as warranted and deemed necessary by the respective State Soil Scientist. The Soil Data Warehouse will archive all older versions of datasets that occur after December 2003. These older versions will be available if needed by special request. Only the most current version of soil data is available directly via the Soil Data Mart and Geospatial Data Warehouse. A subscription service is provided on the Soil Data Mart to inform customers when updated datasets have been posted. Accompanying metadata files should be updated to reflect changes made. If SCA offices choose to archive older versions of datasets, they should be archived in separate folders and files named accordingly. It is recommended that archived files be renamed by adding the av*x* extension to standard file names (av*x* stands for archived version 1, 2, etc.).

Changes to data structure of tabular datasets does happen but is will be minimized to the extent possible, both in frequency and nature of change. Known customers and partners will be notified during the proposed change consideration process and before the actual implementation of such changes. A formal process to do this has not been developed at this time.